#344

DTIC FILE COPY

(2)

DTIC
ELECTE
MAR 26 1990
S D
D CS

AD-A219 686

SPECIFICITIES OF GERM LINE ANTIBODIES

Thomas P. Theriault, Gordon S. Rule and Harden M. McConnell

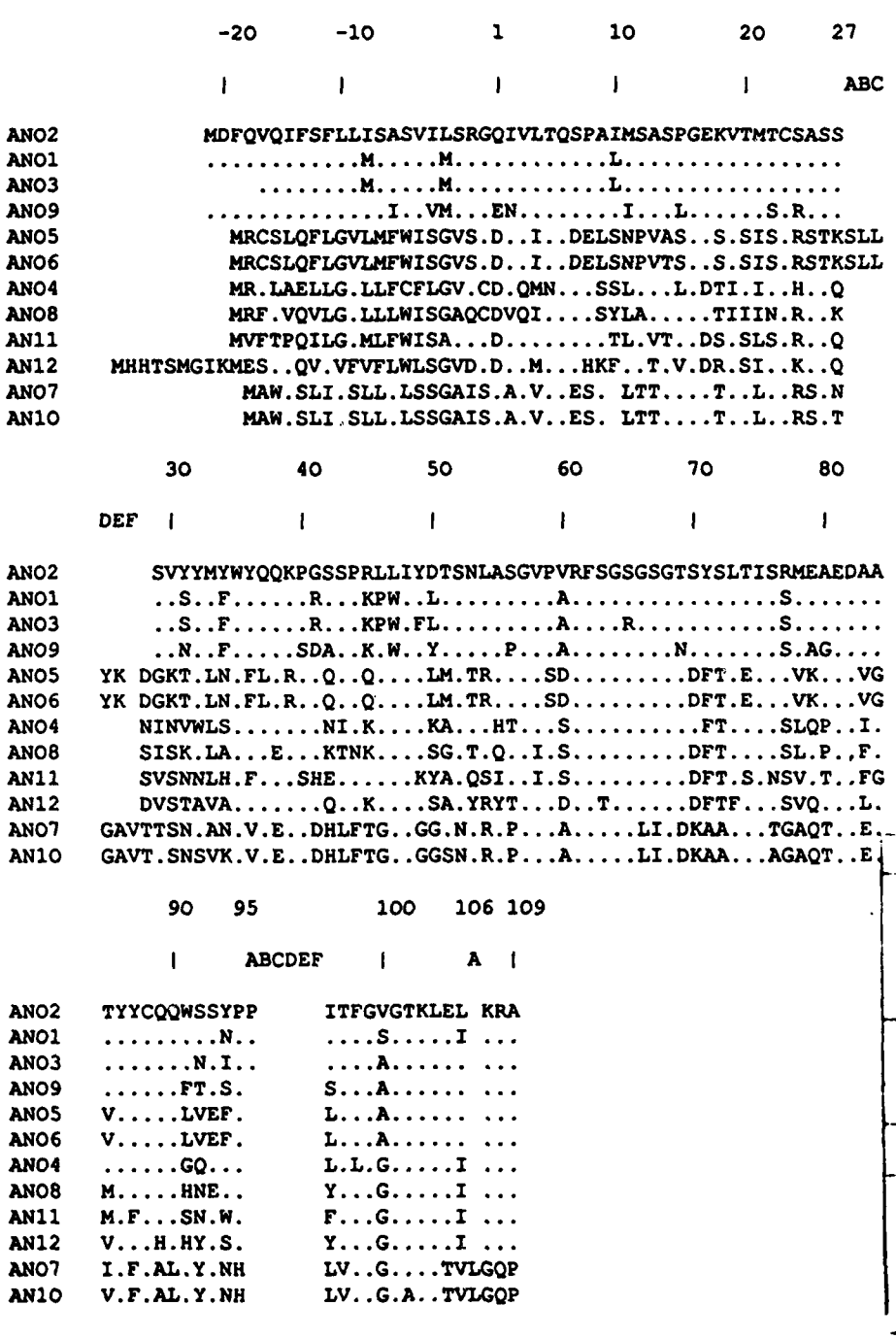
Stauffer Laboratory for Physical Chemistry, Department of Chemistry, Stanford University, Stanford, California 94305, USA

The past few years have seen remarkable progress in research on the structure and function of antibodies. In our own work we have shown that it is possible to obtain extensive significant information on the composition and structure of antibody combining sites using NMR, together with nitroxide spin-label haptens (Anglister et al., 1984a, 1985 and 1987; Frey et al., 1984). This derived information includes the amino acid composition of the combining site region, that is, the number of tyrosines, alanines, etc. that are within ~ 20 Å of the odd electron on the paramagnetic hapten. In antibody molecules there are typically 40-50 amino acids in this combining site region. We have shown that NMR titration data can be used to estimate distances between individual protons on amino acid side chains and the odd electron (Anglister et.al., 1984b; Frey et. al., 1988). These measured distances extend out to about 20 Å and in to distances of the order of 3-5 Å. Shorter distances can sometimes be estimated from nuclear magnetization transfer experiments. The NMR data also provide a powerful and convenient means of obtaining the on-off kinetics of hapten-antibody reactions, using resonance signals from the hapten as well as from the protein.

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

Our NMR studies are being carried out on 12 monoclonal anti-dinitrophenyl nitroxide spin label antibodies that we have prepared and sequenced (cDNA) (Leahy et al., 1988). Theoretical models for the Fab fragments of all of these antibodies have been made in collaboration with Dr. Michael Levitt (unpublished). A major purpose of our current work is to compare these theoretical structure models and the NMR data. An essential requirement for this work is the production of antibody mutants that can be used to obtain rigorous NMR assignments.

Figures 1 and 2 show schematically the sequence similarities of the Fab fragments of the antibodies AN01-AN12. Some pairs of these antibodies have sequences that are very similar in both heavy and light chains (AN01,2,3), or (AN05,6), some pairs have similarities in only the heavy chains (AN04,11,12), and other pairs have similarities in only the light chains (AN07-AN10). Clearly many sequences can accommodate the same hapten. Since germline antibody genes doubtless evolved long before dinitro-phenyl was first synthesized, we have developed an interest in the question of whether the germline antibodies would also recognize the dinitrophenyl nitroxide spin label hapten. If the germline antibodies did not recognize this hapten, then we would conclude that their specificity/affinity was largely developed through somatic mutation. If the germline antibodies did recognize the dinitrophenyl nitroxide spin label hapten with essentially the same specificity/affinity, then we would conclude that the specificity



Accession For
NTIS CRA&I
DTIC TAB
Unannounced
Justification
By
Distribution /
Availability Codes
Dist
A-1

Figure 1. Deduced amino acid sequences of the V regions of the light chains of the anti-DNP-SL monoclonal antibodies AN01-AN12. Numbering system according to Kabat et al. (1987). From Leahy et al. (1988).

of these antibodies was essentially accidental, and we might then speculate on the nature of the primordial antigen and how its structure might be related to dinitrophenyl.

QUALITY INSPECTED 4

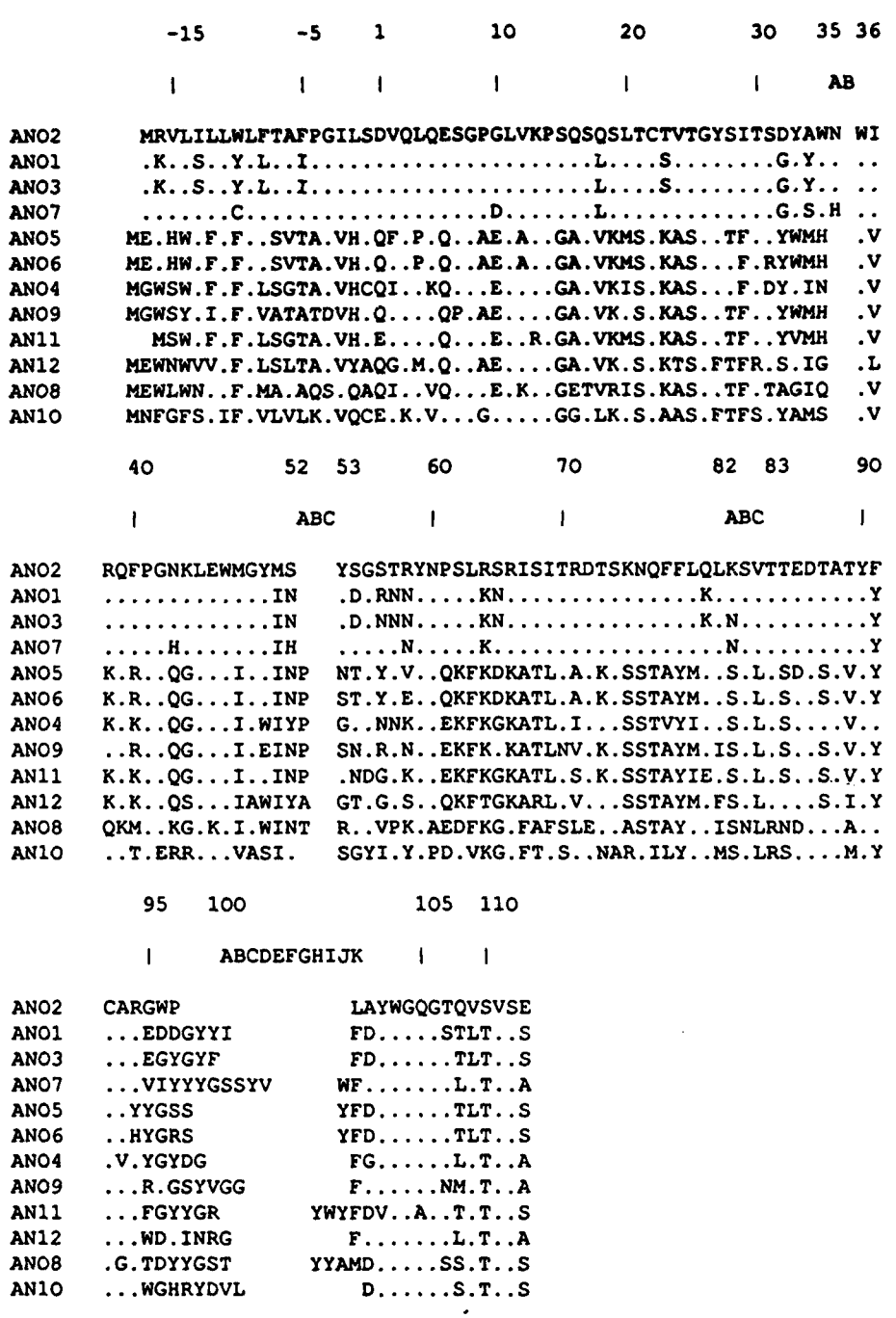


Figure 2. Deduced amino acid sequences of the V regions of the heavy chains of the anti-DNP-SL monoclonal antibodies AN01-AN12. Numbering system according to Kabat et al. (1987). From Leahy et al. (1988).

MATERIALS AND METHODS

Southern Blot Analysis

Balb-c genomic DNA was isolated from the livers of Balb-c mice as described in Maniatis et al. (1982). AN02 genomic DNA was similarly isolated from AN02 antibody producing

hybridoma cells. Each of these DNAs were digested with a series of restriction enzymes and analyzed by the method of Southern (1975). The probes used in the hybridization were prepared from the variable regions of the light and heavy chain coding sequences of an AN02 containing plasmid. The light chain probe was a 1000 base pair fragment made by digesting the AN02 genomic kappa chain clone with BglII and PpuMI. This probe begins approximately 800bp 5' to the variable region. The heavy chain probe was a 600bp fragment isolated as an XbaI and EcoRI fragment from an AN02 heavy chain clone which extends to within 40 base pairs of the end of the variable region. The probes were labelled using Pharmacia's oligo labelling kit with α-32P dCTP (New England Nuclear) as label. The hybridizations were carried out as described in Maniatis (1982) at 42 degrees Celsius with formamide concentrations ranging from 35% to 50%.

Germline gene cloning and sequencing

Balb-c liver DNA was digested with the appropriate restriction enzyme, size selected on an agarose gel and isolated using an International Biotechnology Unidirectional Electroelutor. Libraries were made by ligating these fragments into either lambda ZAP or lambda NM590. Recombinant phage DNA was packaged as described in Maniatis (1982) and the resultant libraries were screened using either the heavy or light chain probe. Positive clones were purified and subcloned into pUC322 from lambda NM590 or excised from lambda ZAP into its pBluescript phagemid. The plasmid clones were mapped with 15 to 35 combinations of single and double restriction digests. The clones were subcloned again into M13 vectors and sequenced by the dideoxy method with α-35S dATP as label (New England Nuclear).

RESULTS

Rearrangement of Balb-c DNA

Genetic rearrangement as indicated by restriction fragment length polymorphism is evidenced by the results of the Southern blot experiments shown in Figures 3 and 4. In the case of the heavy chain for AN02, the DNA sequence of the variable region was known to share 97.6% sequence identity with a heavy chain Balb-c germline sequence named SB32 and with the variable region of a gene isolated from an IgM producing hybridoma LB8 published by Dzierzak et al. (1986). The hybridization of fragments obtained by digestion with the enzymes EcoRI, HindIII, and XbaI shows a pattern consistent with what had been seen previously, namely hybridization to a 2.4 kb EcoRI fragment, an 7.5 kb HindIII fragment, and a 7.0 kb XbaI fragment. The AN02 genomic heavy chain gene is known to be contained on a 4.6 kb EcoRI and a 1.9 kb XbaI fragment (G. Rule unpublished data). These fragments are seen to hybridize as would be expected. Our assumption then was that a Balb-c germline variable region gene which was contained on 2.4 kb EcoRI and 7.0 kb XbaI fragments was rearranged by the processes of recombination and somatic mutation such that it rested on 4.6 kb EcoRI and 1.9 kb XbaI fragments in the AN02 genome. The results for digestion with HindIII are less clear, since the size of the restriction fragment which contains the AN02 genomic heavy chain gene is not known. The patterns seen in the double digestion with XbaI/HindIII and XbaI/KpnI, indicate that the single band in the XbaI lane results from more than one germline variable region.

In the case of the light chain a large number of fragments cross hybridize with the AN02 light chain probe indicating that the AN02 kappa variable gene comes from a large family of related germline genes. Most of the hybridizing bands seen in the Balb-c lanes are matched by equivalent bands in AN02. The major difference is that a 3.7 kb HindIII fragment hybridizes with approximately twice as much intensity in Balb-c as in AN02. This pattern would be expected if only one of the allelic light chain containing chromosomes were recombined during

the differentiation of the AN02 producing B lymphocyte and if the rearranged germline variable gene were situated at the 3' end of the family of cross reacting variable regions. Further evidence that the AN02 light chain germline is contained on a 3.7 kb HindIII fragment comes from performing the hybridization under more stringent conditions. Only the bands at 3.5, 3.7, and 3.9 kb appear, and the 3.7 kb band is the most intense.

Figure 3. Southern blot results using the light chain probe. AN02 hybridoma and Balb-c genomic DNA were digested with HINDIII and EcoRI. Fragmant sizes were internally referenced by including HINDIII digested lambda DNA in the gel and as a radiolabelled probe. Shown are experiments in which the hybridizations were carried out in 50% and 35% formamide.

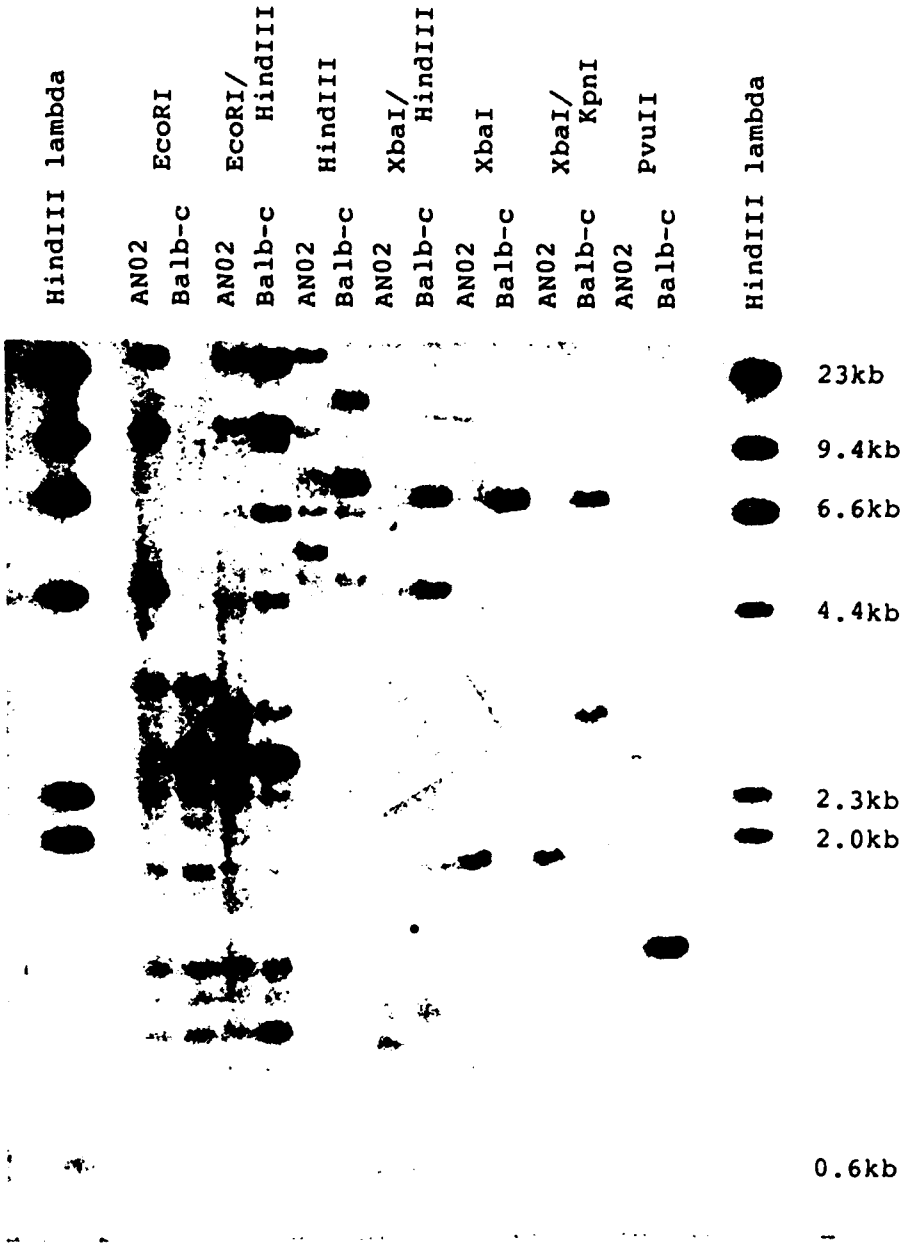


Figure 4. Southern blot results using the heavy chain probe. AN02 hybridoma and Balb-c genomic DNA were digested with the same enzymes. Fragmant sizes were internally referenced by including HINDIII digested lambda DNA in the gel and as a radiolabelled probe.

Cloning and sequencing of germline genes

To determine whether the SB32 gene was indeed the germline gene for AN02 or if a gene with even more sequence identity exists, the 7.0 kb XbaI as well as the 7.5 kb HindIII fragments were ligated into lambda ZAP and lambda NM590 respectively. Multiple positive clones were purified and sequenced from both libraries and in each case a germline gene was obtained that was more than 99% sequence identity to the heavy chain variable region of the

antibody AN07 and only 92% sequence identity to AN02. Since the antibody AN07 has a lambda type light chain, the complete germline configuration of AN07 could be deduced. The inability to find the SB32 germline gene or any other AN02-like gene in the XbaI or HindIII libraries upon sequencing 14 independent clones indicated that there may have been a systematic problem with the vectors or hosts used in the cloning process. To better understand the clonability of these variable genes, as well as to obtain partial sequence data for other germline variable regions, a library was constructed which contained EcoRI/XbaI Balb-c DNA fragments of approximately 600 base pairs in length. Two AN02-like germline genes were sequenced from this library in the five clones which have been purified. One of these was the SB32 germline, the other was a gene that was only 87% similar to AN02.

For the light chain germline gene libraries were made in lambda NM590 with HindIII fragments of approximately 3.5, 3.7, and 3.9 kb. Positive clones were purified and sequenced in each case. Two light chain germline genes were found on the 3.5 kb fragments which were 94% similar to the AN02 light chain, one was found on the 3.9 kb fragment that was 96% similar, and one was found on the 3.7 kb fragment which had more than 98% sequence identity. This is assumed to be the germline gene for the light chain variable region of AN02.

Construction of antibody genes in the recombined germline state.

In order to investigate the binding properties of the IgM class antibody which was found on the surface of the lymphocyte clone that eventually mutated into AN02 or AN07, it is necessary to first make genetic constructs which reflect the sequence of this germline antibody and then express these constructs in a suitable system. Taking the variable regions described above as the true germline variable genes, the remainder of the recombined antibody genes can be found in the germline joining and diversity sequences which have been well characterized (Kabat et al. 1987). For AN07 the diversity segment DFL16.1 and the joining segments JH3 and JL1 were used. For AN02 the diversity segment was too short to determine unambiguously while the joining segments JH3 and JK5 were used. The inferred protein sequence of the heavy and light chain cDNA clones for AN02 and AN07 together with that of the corresponding germline variable regions, diversity segments and joining regions recombined in the proper reading frame are shown in Figures 5 and 6. In each case more than 97% of the protein residues for each

```
         1        10        20     27         30        40        50
         |        |         |      ABCDEF     |         |         |
AN02     QIVLTQSPAIMSASPGEKVTMTCSASS       SVYYMYWYQQKPGSSPRLLIYDTSNL
K5a      ...........................       ..S.......................

AN07     QAVVTQESA LTTSPGETVTLTCRSSN   GAVTTSNYANWVQEKPDHLFTGLIGGTNNR
Lva      ......... ................T   ..............................

              60        70        80        90   95          100  106
              |         |         |         |    ABCDEF      |     A
AN02     ASGVPVRFSGSGSGTSYSLTISRMEAEDAATYYCQQWSSYPP      ITFGVGTKLEL KR
K5a      ..........................................      ....A...... ..

AN07     APGVPARFSGSLIGDKAALTITGAQTEDEAIYFCALWYSNH       LVFGGGTKLTVLG
Lva      .........................................       .............
```

Figure 5. Sequences of AN02 and AN07 light chains compared with their germline gene sequences. Numbering as in Kabat et al. (1987).

```
        1         10        20        30   35     40          52  53      60
        |         |         |         |    AB     |           ABC       |
AN02    DVQLQESGPGLVKPSQSQSLTCTVTGYSITSDYAWN WIRQFPGNKLEWMGYMS   YSGSTRYNP
Gan2    ................L.................. ..............I.   .....S...

AN07    DVQLQESGPDLVKPSQSLSLTCTVTGYSITSGYSWH WIRQFPGHKLEWMGYIH   YSGSTNYNP
G12a    ..................................... .......N.........   .........


                70          82  83     90   95   100             105  110
                |            ABC       |    |    ABCDEFGHIJK     |    |
AN02    SLRSRISITRDTSKNQFFLQLKSVTTEDTATYFCARGWP                LAYWGQGTQVSVSE
Gan2    ..K.................N.........Y...XXX                 X.......L.T..A

AN07    SLKSRISITRDTSKNQFFLQLNSVTTEDTATYYCARVIYYYGSSYV      WFAYWGQGTLVTVSA
G12a    ..................................XXX......XX      ...............
```

Figure 6. Sequences of AN02 and AN07 heavy chains compared with their germline gene sequences. Numbering as in Kabat et al. (1987).

antibody can be assigned to their germline states. The ambiguities that exist occur at the junctions between the variable regions, joining, and diversity segments. At splice junctions it is impossible to determine which germline splice partner donated the germline sequence since mutation after recombination could convert one sequence into the other. Compounding this problem in the heavy chains is the mechanism of nucleotide insertion which can place random nucleotides into the V-D and D-J junctions during recombination. However, given that the observed mutation rate in antibody genes during somatic diversification is approximately 0.5% (Gearhart et al. 1983), the small set of ambiguous nucleotides is highly likely to have been the same in the germline as in the mature antibody gene. In the worst case, AN02 with 11 ambiguous nucleotides, the probability is still approximately 95% that the inferred protein sequence for all of these codons correctly represent the germline.

Future studies

Site specific mutatgenesis has already been used to change the AN02 light chain cDNA clone into the germline gene sequence defined above. Progress is currently underway to construct the other germline genes from the cDNA clones by either site directed mutagenesis or where suitable restriction sites exist, by replacement with the germline DNA itself. These constructs will then be used in an expression system from which mutant protein can be obtained. Studies will then center on characterizing these proteins in terms of their binding affinities and kinetics for various DNP derivatives.

DISCUSSION

The results of the germline sequencing for AN07 indicate that very little somatic mutation occurred for this antibody. Only two residue changes exist between the known germline genes and the cDNA clones. One of these is removed from the complementary determining regions and is not considered to be a factor. The other occurs in CDR 1 of the light chain. Modelling studies suggest that this residue is removed from the binding pocket and will not be in contact

with the bound hapten. Thus we expect the germline antibody and AN07 to have the same affinity for the DNP haptens.

The AN02 germline has a total of 11 residue differences from the sequence of AN02, however only 3 of these occur in the hypervariable loops. Although two of these occur in residues of the second hypervariable loop of the heavy chain they are well removed from our model of the binding pocket and are unlikely to have any significant effect on the binding of the hapten. The residue change in CDR 2 of the light chain, Ser in the germline to Tyr in AN02 at position 31, can also be considered unimportant since NMR studies of tyrosines have implicated only one tyrosine, Tyr34L, as being in contact with the hapten (Rule et al., in preparation). We believe that the antibody in its germline state will have a similar affinity and show similar kinetics to the mature antibody as far as the DNP haptens are concerned.

The plausible assumption that the germline antibodies bind DNP haptens leads us to the conclusion that the DNP specificities of these antibodies are essentially accidental since the germline genes could have only experienced evolutionary pressure in response to naturally occurring antigens. The question then arises as to what the primordial antigen might have been. For example, on the molecular scale, was it big or small? Our NMR studies of the antibodies AN01, AN02, and AN03 show a remarkably large concentration of aromatic residues, especially tyrosines, in the combining site region. In our most studied molecule, AN02, there are 8-10 tyrosine residues in this region (Anglister et al., 1984), many more than could possibly all be in contact with the hapten. Further, as will be discussed elsewhere (Rule et al., in preparation), the sharp proton resonance signals from these residues, with one notable exception, show almost no change on hapten binding, leading us to believe they play no functional role in the binding of DNP haptens. Kinetics studies of antibody-hapten binding also support the view that no significant change in protein structure takes place on hapten binding. We therefore suggest that the primordial antigen recognized by these germline genes was large, with a size of the order of a protein antigen. Studies of other antibody-protein complexes (Amit et al., 1986; Sheriff et al., 1987) typically show that the antibody recognition area is large, on the order of 700 square angstroms of interface. Furthermore, AN02 has been shown to be a cryoglobulin (Theriault et al., in preparation), indicating a possible role for the binding site aromatic residues in idiotype-antiidiotype regulation, the idiotype here being self protein. These ideas will be developed in fuller detail elsewhere.

ACKNOWLEDGEMENTS

We are pleased to thank Daniel Leahy and Dan Denney for providing assistance and guidance in molecular cloning. This work was supported by Office of Naval Research contract N00014-86-k-0388 and American Cancer Society grant NP-631.

REFERENCES

Amit, A. G., Mariuzza, R. A., Phillips, E. V. and Poljak, R. J., 1986, Three-dimentional structure of an antigen-antibody complex at 2.8 Å resolution, *Science* 233:747.

Anglister, J., Frey, T. and McConnell, H. M., 1984, Magnetic resonance of a monoclonal anti-spin-label antibody, *Biochemistry* 23:1138.

Anglister, J., Frey, T. and McConnell, H. M., 1984, Distances of tyrosine residues from a spin-label hapten in the combining site of a specific monoclonal antibody, *Biochemistry* 23:5372.

Anglister, J., Bond, B. W., Frey, T, Leahy, D. J., Levitt, M., McConnell, H. M., Rule, G. S., Tomasello, J. and Whittaker, M. M., 1987, Contribution of tryptophan residues to the combining site of a monoclonal anti-dinitrophenyl spin-label antibody, *Biochemistry* 26:6058.

Anglister, J., Frey, T. and McConnell, H. M., 1985, NMR technique for assessing contributions of heavy and light chains to an antibody combining site, *Nature* 315.6014:65.

Dzierzak, E. A., Janeway, C. A., Jr., Richard, N. and Bothwell, A., 1986, Molecular characterization of antibodies bearing Id-460, *J. Immunol.* 136.5:1864.

Frey, T., Anglister, J. and McConnell, H. M., 1984, Nonaromatic amino acids in the combining site of a monoclonal anti-spin-label antibody, *Biochemistry* 23:6470.

Frey, T., Anglister, J. and McConnell, H. M., 1988, Line-shape analysis of NMR difference spectra of an anti-spin-label antibody, *Biochemistry* 27:5161.

Gearhart, P. J. and Brogenhagen, D. F., 1983, Clusters of point mutations are found exclusively around rearranged antibody variable genes, *Proc. Natl. Acad. Sci. USA* 80:3439.

Leahy, D. J., Rule, G. S., Whittaker, M. M. and McConnell, H. M., 1988, Sequences of 12 monoclonal anti-dinitrophenyl spin-label antibodies for NMR studies, *Proc. Natl. Acad. Sci. USA* 85:3661.

Kabat, E. A., Wu, T. T., Reid-Miller, M., Perry, H. M. and Gottesman, K. S., 1987, "Sequences of Proteins of Immunological Interest", Natl. Inst. of Health, Bethesda, MD.

Maniatis, T., Fritsch, E. F. and Sambrook, J., 1982, "Molecular Cloning : A Laboratory Manual", Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sheriff, S., Silverton, E. W., Padlan, E. A., Cohen, G. H., Smith-Gill, S. J., Finzel, B.C. and Davies, D. R., 1987, Three-dimensional structure of an antibody-antigen complex, *Proc. Natl. Acad. Sci. USA* 84:8075.

Southern, E., 1975, Detection of specific sequences among DNA fragments separated by gel electrophoresis, *J. Mol. Biol.* 98:503.